

OM nucleic - nucleic search, using sw model

Run on: July 7, 2004, 17:46:47 ; Search time 3204.75 seconds  
(without alignments)  
16080.756 Million cell updates/sec

Title: US-10-018-786-3  
Perfect score: 1189  
Sequence: 1 aaattcgaagggtgatgc.....gcttgtgaccagtggaag 1189

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl:  
1: gb\_ba:  
2: gb\_htg:  
3: gb\_in:  
4: gb\_ov:  
5: gb\_ov:  
6: gb\_pat:  
7: gb\_ph:  
8: gb\_pl:  
9: gb\_pr:  
10: gb\_ro:  
11: gb\_sts:  
12: gb\_sy:  
13: gb\_un:  
14: gb\_vi:  
15: em\_ba:  
16: em\_fun:  
17: em\_hum:  
18: em\_in:  
19: em\_mu:  
20: em\_ov:  
21: em\_or:  
22: em\_ov:  
23: em\_pat:  
24: em\_ph:  
25: em\_pl:  
26: em\_ro:  
27: em\_sts:

28: em\_un:  
29: em\_vi:  
30: em\_htg\_hum:  
31: em\_htg\_inv:  
32: em\_htg\_other:  
33: em\_htg\_mus:  
34: em\_htg\_pln:  
35: em\_htg\_rnd:  
36: em\_htg\_mam:  
37: em\_htg\_vrt:  
38: em\_sy:  
39: em\_htg\_hum:  
40: em\_htg\_mus:  
41: em\_htg\_other:

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	1189	100.0	1189	6	AX061802	AX061802 Sequence
c 2	1183.2	99.5	11629	1	AE012222	AE012222 Xanthomon
3	944.8	79.5	1938	1	XANHRPC2A	M99176 Xanthomonas
c 4	944.8	79.5	5071	1	AF320050	AF320050 Xanthomon
5	944.8	79.5	6454	1	AY139029S2	AY139029 Xanthomon
6	944.8	79.5	19304	1	AB045311	AB045311 Xanthomon
7	944.8	79.5	23514	1	AY205561	AY205561 Xanthomon
8	943.2	79.3	5083	1	AB040134	AB040134 Xanthomon
9	943.2	79.3	100258	1	AY055110	AY055110 Xanthomon
c 10	940	79.1	13551	1	AE011666	AE011666 Xanthomon
11	938.4	78.9	29206	1	AF495777	AF495777 Xanthomon
12	932	78.4	6125	1	AF160974	AF160974 Xanthomon
13	374.8	31.5	2073	1	PSEHRPAA	M99633 Pseudomonas
c 14	374.8	31.5	23407	1	RS0245811	AJ245811 Ralstonia
c 15	374.8	31.5	197050	1	AL646081	AL646081 Ralstonia
16	316.6	26.6	3177	1	AB053454	AB053454 Acidovor
17	314	26.4	2920	1	AB053455	AB053455 Burkholde
18	270.2	22.7	29814	1	AF074878	AF074878 Burkholde
19	164.2	13.8	15231	1	AY166598	AY166598 Burkholde
20	164.2	13.8	34600	1	AY028431	AY028431 Burkholde
21	144.6	12.2	4889	1	AF010150	AF010150 Pseudomon
22	144.6	12.2	10052	1	AE004597	AE004597 Pseudomon
23	142.8	12.0	6899	1	AY463491	AY463491 Pseudomon
c 24	142.8	12.0	300917	1	AE016919	AE016919 Chromobac
25	142	11.9	2100	6	BD269171	BD269171 Vaccine.
26	142	11.9	2100	6	AX028406	AX028406 Sequence
c 27	142	11.9	35026	1	BPHYPLCRD	Y13383 Bordetella
28	142	11.9	348934	1	BX640417	BX640417 Bordetell
c 29	140.6	11.8	301870	1	AF005078	AF005078 Vibrio pa
30	137.6	11.6	5602	1	ASA458292	AJ458292 Aeromonas
31	137.6	11.6	5678	6	AX566769	AX566769 Sequence
32	133.6	11.2	10512	1	AY169276	AY169276 Erwinia c
c 33	132.4	11.1	344321	1	BX640429	BX640429 Bordetell

34	131.4	11.1	52498	1	AF232004	AF232004 Pseudomon
c 35	131.4	11.1	311249	1	AE016860	AE016860 Pseudomon
36	131.2	11.0	300917	1	AE016919	AE016919 Chromobac
37	130.8	11.0	110000	1	FEU80928_2	Continuation (3 of
c 38	130.8	11.0	348624	1	EX640441	EX640441 Bordetell
39	128.8	10.8	3603	1	PSEHRFIA	L11582 Pseudomonas
c 40	128.2	10.8	300450	1	AP005960	AP005960 Bradyrhiz
41	119.8	10.1	28666	1	AE293288	AE293288 Pectobact
c 42	115	9.7	11518	1	ERWHRPI	L25928 Erwinia amy
c 43	111.6	9.4	215050	1	AL646084	AL646084 Ralstonia
c 44	111.4	9.4	349944	1	EX571871	EX571871 Photorhab
c 45	111.4	9.4	349940	6	AX770910	AX770910 Sequence

  

ALIGNMENTS									
RESULT 1	AX061802	AX061802	1189 bp	DNA	linear	PAT 24-JAN-2001			
LOCUS	Sequence 3 from Patent WO0078967.								
DEFINITION	AX061802								
ACCESSION	AX061802								
VERSION	AX061802.1	GI:12539882							
KEYWORDS	Xanthomonas campestris								
ORGANISM	Xanthomonas campestris								
	Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;								
	Xanthomonadaceae; Xanthomonas.								
REFERENCE 1	Pierrat, J., Simon, J.L. and Chevallereau, P.								
AUTHORS	Avirulent xanthomonas-campestris strains producing xanthan								
TITLE	Patent: WO 078967-A 3 28-DEC-2000;								
JOURNAL	RHODIA CHIMIE (FR)								
FEATURES	Location/Qualifiers								
source	1..1189								
	/organism="Xanthomonas campestris"								
	/mol_type="unassigned DNA"								
	/db_xref="taxon:339"								
ORIGIN									
Query Match	100.0%;	Score 1189;	DB 6;	Length 1189;					
Best Local Similarity	100.0%;	Pred. No. 1.1e-172;							
Matches 1189;	Conservative	0;	Mismatches	0;	Indels	0; Gaps			
QY	1	AAATTCGTCAAGGTCATCGATCGCGGCTTGGTGATCACCATGGTCAACATCTTGGCC	60						
Db	1	AAATTCGTCAAGGTCATCGATCGCGGCTTGGTGATCACCATGGTCAACATCTTGGCC	60						
QY	61	GGCATTCGTAGGCGTGACCTACACGCGCATGAGCGCGGCGAGCGCCGACACCGCTTT	120						
Db	61	GGCATTCGTAGGCGTGACCTACACGCGCATGAGCGCGGCGAGCGCCGACACCGCTTT	120						
QY	121	CGCATTCGTGGTAGGCGATGCGATGGTGTCGACATCGCTGGTGTGATCTGGTG	180						
Db	121	CGCATTCGTGGTAGGCGATGCGATGGTGTCGACATCGCTGGTGTGATCTGGTG	180						
QY	181	CGCGCGCGGCGTCAATCATCCCGCGGTGCGCAACGAGATGAACCAAGATCAGCTCGCTC	240						

Db 1021 GCGCGCTATCTTGGCACACCGGACCGGACCGGACGCTGCTGGGTGATGCTC 1080

QY 1081 GACACCGCGTGGAAACAGTGTATCGGAGTCGATCGGACACACCGCGCGCAATTC 1140  
 |||||

Db 1081 GACACCGCGTGGAAACAGTGTATCGGAGTCGATCGGACACACCGCGCGCAATTC 1140  
 |||||

QY 1141 CTGGCGGTGCCACCGGACGCGCAATCAGCTTGTGACACAGGTGGAAA 1189  
 |||||

Db 1141 CTGGCGGTGCCACCGGACGCGCAATCAGCTTGTGACACAGGTGGAAA 1189  
 |||||

RESULT 2  
 AE012222/c 11629 bp DNA linear ECT 23-MAY-2002  
 LOCUS Xanthomonas campestris pv. campestris str. ATCC 33913, section 130  
 DEFINITION of 460 of the complete genome.  
 ACCESSION AE012222 AE008922  
 VERSION AE012222.1 GI:21112273  
 KEYWORDS  
 SOURCE Xanthomonas campestris pv. campestris str. ATCC 33913  
 ORGANISM Xanthomonas campestris pv. campestris str. ATCC 33913  
 Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;  
 Xanthomonadaceae; Xanthomonas.  
 REFERENCE 1 (bases 1 to 11629)  
 AUTHORS da Silva,A.C.R., Ferro,J.A., Reinach,F.C., Farah,C.S., Furlan,L.R.,  
 Quaggio,R.B., Monteiro-Vitorello,C.B., Van Sluys,M.A., Almeida  
 Jr.,N.F., Alves,L.M.C., do Amaral,A.M., Bertolini,M.C.,  
 Camargo,L.E.A., Camarotte,G., Cannavan,F., Cardozo,J.,  
 Chamergo,F., Ciapina,L.P., Cicarelli,R.M.B., Coutinho,L.L.,  
 Cursino-Santos,J.R., El-Dorri,H., Faria,J.B., Ferreira,A.J.S.,  
 Ferreira,R.C.C., Ferro,M.I.T., Fornighieri,E.F., Franco,M.C.,  
 Greggio,C.C., Gruber,A., Katsuyama,A.M., Kishi,L.T., Leite  
 Jr.,R.P., Lemos,E.G.M., Lemos,M.V.F., Locali,E.C., Machado,M.A.,  
 Madalira,A.M.B.N., Martinez-Rossi,N.M., Martins,E.C., Meidanis,J.,  
 Menck,C.F.M., Miyaki,C.Y., Moon,D.H., Moreira,L.M., Novo,M.T.M.,  
 Okura,V.K., Oliveira,M.C., Oliveira,V.R., Pereira Jr.,H.A.,  
 Rossi,A., Sena,J.A.D., Silva,C., de Souza,R.F., Spindola,L.A.F.,  
 Takita,M.A., Tamura,R.E., Teixeira,E.C., Tezza,R.I.D., Trindade dos  
 Santos,M., Truffi,D., Tsai,S.M., White,F.F., Setubal,J.C. and  
 Kitajima,J.P.  
 Comparison of the genomes of two Xanthomonas pathogens with  
 differing host specificities  
 JOURNAL Nature 417 (6887), 459-463 (2002)  
 MEDLINE 22022145  
 PUBMED 12024217  
 REFERENCE 2 (bases 1 to 11629)  
 AUTHORS da Silva,A.C.R., Ferro,J.A., Reinach,F.C., Farah,C.S., Furlan,L.R.,  
 Quaggio,R.B., Monteiro-Vitorello,C.B., Van Sluys,M.A., Almeida  
 Jr.,N.F., Alves,L.M.C., do Amaral,A.M., Bertolini,M.C.,  
 Camargo,L.E.A., Camarotte,G., Cannavan,F., Cardozo,J.,  
 Chamergo,F., Ciapina,L.P., Cicarelli,R.M.B., Coutinho,L.L.,  
 Cursino-Santos,J.R., El-Dorri,H., Faria,J.B., Ferreira,A.J.S.,  
 Ferreira,R.C.C., Ferro,M.I.T., Fornighieri,E.F., Franco,M.C.,  
 Greggio,C.C., Gruber,A., Katsuyama,A.M., Kishi,L.T., Leite  
 Jr.,R.P., Lemos,E.G.M., Lemos,M.V.F., Locali,E.C., Machado,M.A.,  
 Madalira,A.M.B.N., Martinez-Rossi,N.M., Martins,E.C., Meidanis,J.,  
 Menck,C.F.M., Miyaki,C.Y., Moon,D.H., Moreira,L.M., Novo,M.T.M.,

Okura,V.K., Oliveira,M.C., Oliveira,V.R., Pereira Jr.,H.A.,  
 Rossi,A., Sena,J.A.D., Silva,C., de Souza,R.F., Spindola,L.A.F.,  
 Takita,M.A., Tamura,R.E., Teixeira,E.C., Tezza,R.I.D., Trindade dos  
 Santos,M., Truffi,D., Tsai,S.M., White,F.F., Setubal,J.C. and  
 Kitajima,J.P.  
 Direct Submission  
 Submitted (28-NOV-2001) Departamento de Bioquímica, Universidade de  
 Sao Paulo, Av. Prof. Lineu Prestes 746, Sao Paulo, SP 05508-900,  
 Brazil  
 Location/Qualifiers  
 1. 11629  
 /organism="Xanthomonas campestris pv. campestris str. ATCC  
 33913"  
 /mol\_type="genomic DNA"  
 /strain="ATCC 33913"  
 /db\_xref="ATCC:33913"  
 /db\_xref="taxon:190485"  
 /note="pathovar: campestris"  
 complement(97. .708)  
 /gene="hpap"  
 /note="synonym: KCCL228"  
 complement(97. .708)  
 /gene="hpap"  
 /note="identified by sequence similarity; putative; ORF  
 located using Blastx/Glimmer/Genemark"  
 /codon\_start=1  
 /transl\_table=11  
 /product="Hpap protein"  
 /protein\_id="AAM40526.1"  
 /db\_xref="GI:21112274"  
 /translation="MRILPIAGPIQPHAAAPAPATQVSMQLRRSLVSLAGLPD  
 TVPEDDDPTLAENVDAASDDHGPAALPTAPCLDDGRRQIARGDDGDLGRQIA  
 TEMRTQRAQLAIHIALVAFNCPVPGAGTWEALNHIDLPQTTLRLRSPD  
 QLSRLDSSSPETREVLNKRRLDTALKAALSDTLQISIEVV"  
 complement(732. .2654)  
 /gene="hrcv"  
 /note="synonym: KCCL229"  
 complement(732. .2654)  
 /gene="hrcv"  
 /note="identified by sequence similarity; putative; ORF  
 located using Blastx/Glimmer/Genemark"  
 /codon\_start=1  
 /transl\_table=11  
 /product="Hrcv"  
 /protein\_id="AAM40527.1"  
 /db\_xref="GI:21112275"  
 /translation="MRVTVFYATGEVAALVAVIGLMPLPTPLITLLGINIT  
 LSVLLWMTMVPDSISLSPSILLETLLRLSINTASTKSLHAEAGHIESEGE  
 LVGGLNVLGLIITTVQIVIVKGSERVAENVGARFTLDMGKMSIDADLRGG  
 NTADEARRRRLAIESQLRGMDGKMFVKGDAIGLVITMNLGAVVGVVTHG  
 MSAGEANFAILSVGDAMVSGIASSLISVAAGVMTNVANETKISSLGIOGRQL  
 TSNARALMAASVLLACFAFVPGFALLILAAAVGAGGTYIWRKQRTSGSDQALP  
 STSRKAGKDAPHIRKSAPEFASPLSNRISPOLAARLDPAALDAIESRRQLVLLG  
 LPFGTAIWQSESLQGYEVLIHDVPETRSAISDTADMDKALAAQALAPAHARHLF  
 VGIQETQWMLQGVDPGLVAENVKMPAQRIADVLRLLREIPVENTIKSILSLV  
 VWGPKEDLLMTETVRCDLGRYLAHTATAGTGQLPAVMDLHVAEQILRQISRAITAG  
 NFLALPEQANQLVEQVERIVEDQARHPLAVVASHMDVRYVRREIARLNLEVISYFQ

TITLE  
 JOURNAL

FEATURES  
 source

gene

CDS

gene

CDS

ELGEVQLQPIGRVWA"  
 complement(2663. .3736)  
 /gene="hrcu"  
 /note="synonym: XCC1230"  
 complement(2663. .3736)  
 /gene="hrcu"  
 /note="identified by sequence similarity; putative; ORF  
 located using Blastx/Glimmer/Genemark"  
 /codon\_start=1  
 /transl\_table=11  
 /product="HrcU protein"  
 /protein\_id="AA040528.1"  
 /db\_xref="GI:21112276"  
 /translation="MSDEKTEKPKKLODARRDGEVPISPDTAAAVLAAALVWL  
 AGSYFEHLRALSIGFDTINTDATALHRLGRIGIOGVLLILPEVTAACLAAGLGR  
 TFVQGLNASKVPIEDSLNPNVGNKLSLSLNLKLGKAAVIGVWVYGLR  
 ALMPTIIGLQPPADIAQIGWRLGILCALAVFVLVGAADMSVQWMLFIRDKMS  
 KDQREKSESGDEPKGKKEFAKELVGFDFRERAKAKVMVNPETHYAVALAYEP  
 DGGFLPQVAVGVDEGALLRAYAHNQIPIVANPLARALHELGEAVPESLFTV  
 AVLRLVDELGRDNDESGPLPC"  
 3949. .4404  
 /gene="hrpB1"  
 /note="synonym: XCC1231"  
 3949. .4404  
 /gene="hrpB1"  
 /note="identified by sequence similarity; putative; ORF  
 located using Blastx/Glimmer/Genemark"  
 /codon\_start=1  
 /transl\_table=11  
 /product="HrpB1 protein"  
 /protein\_id="AA040529.1"  
 /db\_xref="GI:21112277"  
 /translation="MEKIECPGVSVGLIELITVGLTHEKIEEAAVLAARVLRPEL  
 KALDTFDAMIARKGSYLEGALLRELDAGSEPLCKALYACCLFVAGDPWHGIAE  
 GLTEEDADADAVGLVKALSGRYTPAPLEAAESAAPMDEVNAQYLRA"  
 4436. .4828  
 /gene="hrpB2"  
 /note="synonym: XCC1232"  
 4436. .4828  
 /gene="hrpB2"  
 /note="identified by sequence similarity; putative; ORF  
 located using Blastx/Glimmer/Genemark"  
 /codon\_start=1  
 /transl\_table=11  
 /product="HrpB2 protein"  
 /protein\_id="AA040530.1"  
 /db\_xref="GI:21112278"  
 /translation="MALIPVPVITASTGLSTPAVSPTAAPNQSLVNRFOALMQSSCP  
 LPPAMORVCEPSMISRWDQNDGVRAMTEHYDAFGANAPKMGLEETAAEKNKLINEL  
 TWGFFNLVSVLSAQSGKNAVQLFNQ"  
 4830. .5594  
 /gene="hrcJ"  
 /note="synonym: XCC1233"  
 4830. .5594  
 /gene="hrcJ"  
 /note="identified by sequence similarity; putative; ORF  
 located using Blastx/Glimmer/Genemark"  
 /codon\_start=1  
 /transl\_table=11  
 /product="HrcJ protein"  
 /protein\_id="AA040531.1"  
 /db\_xref="GI:21112279"  
 /translation="MRTLRYLVLLALLSGCDQQLYSLGTENDANDMLAVLLTAGV  
 DAELKLPDDGKTWVAPHDQVAVLNLVTHGPHERHANLGENFKDGLISTPTEE  
 RVREIVSQCLSTSLNIDGVIADQVIVLNNPDLASVSPSSAAVEIKFRVSGDL  
 TSLVPSIKTLNKHVSLEGLIYENVSVLYPGAESDAQFAASAPPAWAPWLVGCALA  
 LCVAVAAAALYVWPSANRWGQWRURALSRRHAG"  
 5604. .6233  
 /gene="hrpB4"  
 /note="synonym: XCC1234"  
 5604. .6233  
 /gene="hrpB4"  
 /note="identified by sequence similarity; putative; ORF  
 located using Blastx/Glimmer/Genemark"  
 /codon\_start=1  
 /transl\_table=11  
 /product="HrpB4 protein"  
 /protein\_id="AA040532.1"  
 /db\_xref="GI:21112280"  
 /translation="MDKARLNTARWLQRTAAVSTLAQSLDAERVAHWLGPVAAALGW  
 ERAAQRNRVITQMSGWASQVQVGLDPLANRLVLPADLLARVLSALFSPALLAR  
 KCIERELSWFEQRVPAVLDVTRCAGTGATAPLPHDADQAWIGDGRWQADGV  
 WFDPCIAKVLALCLPLGAARVVPITAADGASDAFLRAUPTLLPELPCVFG"  
 6218. .6919  
 /gene="hrpB5"  
 /note="synonym: XCC1235"  
 6218. .6919  
 /gene="hrpB5"  
 /note="identified by sequence similarity; putative; ORF  
 located using Blastx/Glimmer/Genemark"  
 /codon\_start=1  
 /transl\_table=11  
 /product="HrpB5 protein"  
 /protein\_id="AA040533.1"  
 /db\_xref="GI:21112281"  
 /translation="MRLWLRSTPEAVGLDCEVIREALACVLELDAQVHARCAQA  
 LADAQTRAAQLIDAAQRAEAILQDAHRAERSARLGYAAGLRQLDANWERNYHAF  
 AAQDAARARERLAEIVAHACQSVLHGHDPAAALYARAAQALDGALEALQVSVHPD  
 ALDDARRAFDAAAAGGWSNPVELCGDITLALGACVCEMTGVTFETDLRDQLARLRY

Query Match 99.5%; Score 1183.2; DB 1; Length 11629;  
 Best Local Similarity 99.7%; Pred. No. 5.1e-172; Indels 0; Gaps 0;  
 Matches 1185; Conservative 0; Mismatches 3;  
 QY 1 AAATTCGTCAAGGCGGATCGCATCGCGGCGCTGGTGATCACCATGTCATCTTGCC 60  
 Db 2090 AAATTCGTCAAGGCGGATCGCATCGCGGCGCTGGTGATCACCATGTCATCTTGCC 2031  
 QY 61 GGCAATCGGTAGCGGTGACCTACCGCGCATGAGCGCGCGGCGGCGGCGGCGGCGGCTTT 120  
 Db 2030 GGCAATCGGTAGCGGTGACCTACCGCGCATGAGCGCGGCGGCGGCGGCGGCGGCGGCTTT 1971  
 QY 121 GGCAATCGGTAGCGGTGACCTACCGCATGCGGTGCGCATGCGGTGCTGCTGCTGCTGCTG 180

Db 1130 GTGTGGGACCGAAGGAAAAGGATCTGTGATGCTGACCGAGTATGTGGCTGGATCTC 1071  
Qy 1021 GGGCGCTATCTTGGGCACACCGCGACCGGACCGGACAGCTGCTGGGGTGTGCTC 1080  
Db 1070 GGGCGCTATCTTGGGCACACCGCGACCGGACAGCTGCTGGGGTGTGCTC 1011  
Qy 1081 GACGACGCGTGGAAACAGTTGATCGGCGAGTGGATTCGGGOCACACGCGCGCAATTC 1140  
Db 1010 GACGACGCGTGGAAACAGTTGATCGGCGAGTGGATTCGGGOCACACGCGCGCAATTC 951  
Qy 1141 CTGGCGCTGCCACCGGAGCAGGCGCAATCAGCTTGTGGACGAGTGGAA 1188  
Db 950 CTGGCGCTGCCACCGGAGCAGGCGCAATCAGCTTGTGGACGAGTGGAA 903  
Search completed: July 7, 2004, 23:54:44  
Job time : 3212.75 secs

Db 1970 GCGATCTGTGCGTAGCGGATGCGATGGTGTGCGATGCGCTCGCTGATCTGGGTG 1911  
Qy 181 GGGGCGGCGCTCATGATCACCGCGCTGCGCAAGGAAACCAAGATCAAGCTGCTC 240  
Db 1910 GGGGCGGCGCTCATGATCACCGCGCTGCGCAAGGAAACCAAGATCAAGCTGCTC 1851  
Qy 241 GGGCTCGATCATGCGGCGCGCATCACAGCAAGCGACGCTGCTGATGGCAGCGAGTGTG 300  
Db 1850 GGGCTCGATCATGCGGCGCGCATCACAGCAAGCGACGCTGCTGATGGCAGCGAGTGTG 1791  
Qy 301 CTGCTGGGCTGCTTGGCTTGGTGGCGGATTTCCGGGCGCTGCTGTTCTGCTGCTGGA 360  
Db 1790 CTGCTGGGCTGCTTGGCTTGGTGGCGGATTTCCGGGCGCTGCTGTTCTGCTGCTGGA 1731  
Qy 361 GGGGCGGTGCGGTGGCGGGGCTATACGATCTGGGCGAGCAAGCGGACACACGCGGAGC 420  
Db 1730 GGGGCGGTGCGGTGGCGGGGCTATACGATCTGGGCGAGCAAGCGGACACACGCGGAGC 1671  
Qy 421 GATCAGCGCGCATGCGCATCAACGCGCAAGGTTGCCAAGGCGATGCGCGCGCATC 480  
Db 1670 GATCAGCGCGCATGCGCATCAACGCGCAAGGTTGCCAAGGCGATGCGCGCGCATC 1611  
Qy 481 CGAAGAGCGCGCGGATTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCT 540  
Db 1610 CGAAGAGCGCGCGGATTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCT 1551  
Qy 541 GCACGCTCGA CCGCGCGCTGCTGGATCAGCGCATCGAAAGCGAGCGGCAATGCTC 600  
Db 1550 GCACGCTCGA CCGCGCGCTGCTGGATCAGCGCATCGAAAGCGAGCGGCAATGCTC 1491  
Qy 601 GAGCTGCTGGGATTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCT 660  
Db 1490 GAGCTGCTGGGATTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCT 1431  
Qy 661 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720  
Db 1430 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1371  
Qy 721 GCGACATGCAAGAAAGCGCTGGGCGCAAGCGCATCGCACGCTTGCATGCAAGCGCGCAT 780  
Db 1370 GCGACATGCAAGAAAGCGCTGGGCGCAAGCGCATCGCACGCTTGCATGCAAGCGCGCAT 1311  
Qy 781 CTGCTGCTGCGCATCCAGAGAGCGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 840  
Db 1310 CTGCTGCTGCGCATCCAGAGAGCGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1251  
Qy 841 GGGCTGGTTGAGAGTCAACAGGCGATGGCGCGCGCGCGCATCGCGATGCTGTTGCGG 900  
Db 1250 GGGCTGGTTGAGAGTCAACAGGCGATGGCGCGCGCGCGCATCGCGATGCTGTTGCGG 1191  
Qy 901 CGACTGCTGGAAGAGCGATCCCGGTGGCGCAACATCAAGAGCATCTGGAGAGCTGCTG 960  
Db 1190 CGACTGCTGGAAGAGCGATCCCGGTGGCGCAACATCAAGAGCATCTGGAGAGCTGCTG 1131  
Qy 961 GTGTGGGAGCGAGGAAAGGATCTGCTGATGCTGACCGAGTATGTGGCTGGATCTC 1020

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 7, 2004, 17:42:57 ; Search time 347.5 seconds  
(without alignments)  
14535.578 Million cell updates/sec

Title: US-10-018-786-3  
Perfect score: 1189  
Sequence: 1 aaattcgtcaagggtatgctgc.....gcttgctgaccagggtgaaa 1189

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 3373663 seqs, 2124099041 residues 6747726

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_29Jan04:\*

- 1: geneseqn1980s:\*
- 2: geneseqn1990s:\*
- 3: geneseqn2000s:\*
- 4: geneseqn2001as:\*
- 5: geneseqn2001bs:\*
- 6: geneseqn2002as:\*
- 7: geneseqn2003as:\*
- 8: geneseqn2003bs:\*
- 9: geneseqn2003cs:\*
- 10: geneseqn2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1189	100.0	1189	5	ABL57891 Partial h
c 2	203.6	17.1	1218	7	ACA26203
3	142	11.9	2100	3	AAA64849 Bordetell
c 4	142	11.9	35026	3	AAA64890
5	137.6	11.6	5678	6	ABN86172 A. salmon
6	130.8	11.0	2103	3	AAZ45563 Nucleotid
c 7	130.8	11.0	13500	3	AAZ45562 Nucleotid

ABL57891 standard; DNA; 1189 BP.

ABL57891

AC ABL57891;

XX

DT 11-SEP-2003 (revised)

DT 04-JUL-2002 (first entry)

XX

DE Partial hypersensitive reaction and pathogenicity, hrpC2 gene.

XX

KW Hypersensitive reaction and pathogenicity; hrpC2; exo-polysaccharide;

KW xanthan gum; gene; ds.

XX

ALIGNMENTS

c	8	120.8	10.2	1899	7	ACA26479
c 9	111.4	9.4	2118	7	ACF71579	
c 10	111.4	9.4	110000	7	ACF67367_47	
c 11	103.4	8.7	2181	7	ACF65388_00	
c 12	103.4	8.7	2181	7	ACF65388_00	
c 13	98	8.2	34720	2	AAV30458_5	
c 14	98	8.2	34720	2	AAV30458_5	
c 15	95.2	8.0	2124	7	ABX94896	
c 16	89.4	7.5	110000	2	AAZ01425_03	
c 17	88.8	7.5	849	5	AA570994	
c 18	87.8	7.4	2127	9	ADD43765	
c 19	86.6	7.3	609	5	AA573035	
c 20	85.4	7.2	2704	5	AA573037	
c 21	81	6.8	2058	2	AAQ34562	
c 22	81	6.8	2058	7	ACA19671	
c 23	80.2	6.7	958	6	ABK74562	
c 24	79.4	6.7	2310	5	AA592707	
c 25	72.2	6.1	2082	7	ACF68505	
c 26	72.2	6.1	110000	7	ACF67367_12	
c 27	72.2	6.1	182624	7	ACF65379_03	
c 28	71.8	6.0	2350	3	AAZ61505	
c 29	71.8	6.0	110000	2	AAK91990_03	
c 30	71.8	6.0	273254	3	AAAC1914	
c 31	71.4	6.0	43360	8	ACD19193	
c 32	71.4	6.0	43450	9	ADC00771	
c 33	71.4	6.0	45325	8	ACD19247	
c 34	70.6	5.9	13417	2	AA019224	
c 35	70	5.9	660	6	ABX66374	
c 36	70	5.9	1024	6	ABX66703	
c 37	67.4	5.7	16943	9	ADC01147	
c 38	67.4	5.7	16943	9	ADC01147	
c 39	67.4	5.7	16950	8	ACD19153	
c 40	64.4	5.4	500	6	ABK78960	
c 41	63.6	5.3	2501	2	AA062336	
c 42	62.6	5.3	726	6	ABQ19426	
c 43	62.6	5.3	726	6	ABQ19427	
c 44	62.2	5.2	695	4	AAZ25644	
c 45	62	5.2	915	6	ABK74570	

Ac26479 Prokaryot  
Ac71579 Photobab  
Continuation (48 o  
Ac65388 Photobab  
Ac26011 Prokaryot  
Continuation (6 of  
Continuation (6 of  
Abx94896 P. aurogi  
Continuation (4 of  
Aas70994 DNA encod  
Add43765 Chlamydia  
Aas73035 DNA encod  
Aas73037 DNA encod  
Aaq34562 Sequence  
Aca19671 Prokaryot  
Abk74562 Bacillus  
Aas92707 DNA encod  
Acf68505 Photobab  
Continuation (13 o  
Acf65379 Photobab  
Aaz61505 DNA encod  
Continuation (4 of  
Aac81914 Chlamydia  
Acd19193 E. coli o  
Acd00771 Enterobae  
Acd19247 E. coli o  
Aat09224 Partial s  
Abx66374 Helicobac  
Abx66703 Helicobac  
Aaf25596 H. pylori  
Acd01147 Enterobae  
Acd19153 E. coli o  
Abk78960 Bacillus  
Aat62336 Helicobac  
Abq19426 oligonuel  
Abq19427 oligonuel  
Aaf25644 H. pylori  
Abk74570 Bacillus

OS *Xanthomonas campestris*; pv *vesicatoria*.

XX W0200078967-A1.

XX 28-DEC-2000.

XX 21-JUN-2000; 2000WO-FR001725.

XX 22-JUN-1999; 99FR-00007963.

XX (RHOD ) RHODIA CHIM.

XX Pierrard J, Simon J, Chevallereau P;

XX WPI; 2001-102725/11.

XX New *Xanthomonas campestris* bacteria strains for use in production of exo-  
PT polysaccharides are made non-virulent by inactivation of at least one  
PT virulence gene.

XX Claim 17; Page 25-26; 33pp; French.

XX The present invention relates to new *Xanthomonas campestris* bacteria  
CC strains made non-virulent by inactivation of at least one virulence gene  
CC but which have retained the capacity to produce exo-polysaccharides  
CC (preferably xanthan gum). One such virulence gene deleted to produce the  
CC bacterial strains was the hrpZ gene (Hypersensitive Reaction and  
CC Pathogenicity). The hrp genes are essential for pathogenicity in plants.  
CC The present sequence is a partial sequence of the hrpC2, used in an  
CC example from the invention. (Updated on 11-SEP-2003 to standardise OS  
CC field)

XX Sequence 1189 BP; 226 A; 366 C; 392 G; 205 T; 0 U; 0 Other;

Query Match 100.0%; Score 1189; DB 5; Length 1189;  
Best Local Similarity 100.0%; Pred. No. 8.3e-250;  
Matches 1189; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	AAATTCGTCAAGGGTGATGGATCGCGGCGCTGGTGATCAACATGGTCAACATCTTGGGC	60
Db	1	AAATTCGTCAAGGGTGATGGATCGCGGCGCTGGTGATCAACATGGTCAACATCTTGGGC	60
Qy	61	GGCATCTGTAGGGTGATGACCTACACGGGATGAGCGGGGGGAGGGCGGCAACGGCTTT	120
Db	61	GGCATCTGTAGGGTGATGACCTACACGGGATGAGCGGGGGGAGGGCGGCAACGGCTTT	120
Qy	121	GGATTCCTGTGGTAGGGATGGATGGTGGATGGATGGATGGATGGATGGATGGATGGAT	180
Db	121	GGATTCCTGTGGTAGGGATGGATGGTGGATGGATGGATGGATGGATGGATGGATGGAT	180
Qy	181	GGGGCGGGCGTCATGATCACCGCGGTGCGCAACGAGATGAAACGAGATGAGTGGCTC	240
Db	181	GGGGCGGGCGTCATGATCACCGCGGTGCGCAACGAGATGAAACGAGATGAGTGGCTC	240
Qy	241	GGGCTCGATCGGGCGGAGCTCAACAGCAACGAGTGGCTGGTGGTGGTGGTGGTGGT	300
Db	241	GGGCTCGATCGGGCGGAGCTCAACAGCAACGAGTGGCTGGTGGTGGTGGTGGTGGT	300

Qy	301	CTGCTGGCTGCTTTGGCTTGGTGGGGAATTCGGGCGCTGCTGTTCTGCTGCTGGCA	360
Db	301	CTGCTGGCTGCTTTGGCTTGGTGGGGAATTCGGGCGCTGCTGTTCTGCTGCTGGCA	360
Qy	361	GGGGGGTGGTGGGGGGCTATACGATCTGGGCGAAGCAACGCAACGCGGGAGC	420
Db	361	GGGGGGTGGTGGGGGGCTATACGATCTGGGCGAAGCAACGCAACGCGGGAGC	420
Qy	421	GATCAGCCCGCACTGCCATCAACCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	480
Db	421	GATCAGCCCGCACTGCCATCAACCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	480
Qy	481	GGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	540
Db	481	GGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	540
Qy	541	GCACGGCTCGACCGGGCGCTGCTGGATCAGGCGATCGAAAGCGAGCGAGGAGGAT	600
Db	541	GCACGGCTCGACCGGGCGCTGCTGGATCAGGCGATCGAAAGCGAGCGAGGAGGAT	600
Qy	601	GAGCTGCTGGGATTCGGTTCCCGGGATCGCGATATGSCAGAGCAATCCCTGCAGGGC	660
Db	601	GAGCTGCTGGGATTCGGTTCCCGGGATCGCGATATGSCAGAGCAATCCCTGCAGGGC	660
Qy	661	CTGCAGTACGAAGTGTGTGATCCAGGATGTGCGGAAACCGCGAGCGGCTTGAAGGAT	720
Db	661	CTGCAGTACGAAGTGTGTGATCCAGGATGTGCGGAAACCGCGAGCGGCTTGAAGGAT	720
Qy	721	CGGAGATCGAAGGAGGCTGGCCCAACAGGCGATCGACCGCTTGCATGCAAGCGGCGAT	780
Db	721	CGGAGATCGAAGGAGGCTGGCCCAACAGGCGATCGACCGCTTGCATGCAAGCGGCGAT	780
Qy	781	CTGTTGCTCGGATCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	840
Db	781	CTGTTGCTCGGATCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	840
Qy	841	GGGCTGGTTCAGAGGTCACAAAGGCGATGCGGCGGCGGCGGCGGCGGCGGCGG	900
Db	841	GGGCTGGTTCAGAGGTCACAAAGGCGATGCGGCGGCGGCGGCGGCGGCGGCGG	900
Qy	901	CGACTGCTGGAAGAACGCTATCCCGTGCACCAATCAAGAGCATCTTGGAGAGCTGGT	960
Db	901	CGACTGCTGGAAGAACGCTATCCCGTGCACCAATCAAGAGCATCTTGGAGAGCTGGT	960
Qy	961	GTGTGGGAGACGAGGAGGAGGATCTGCTGATGCTGACCGAGTATGTGGCTGCTC	1020
Db	961	GTGTGGGAGACGAGGAGGAGGATCTGCTGATGCTGACCGAGTATGTGGCTGCTC	1020
Qy	1021	GGCGGCTATCTTGGCAGACCGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGG	1080
Db	1021	GGCGGCTATCTTGGCAGACCGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGG	1080
Qy	1081	GACCACCGCGTGGAAACAGTTGATCCGGCAGTGGATTCGGCCACACCGCGGCAATTC	1140
Db	1081	GACCACCGCGTGGAAACAGTTGATCCGGCAGTGGATTCGGCCACACCGCGGCAATTC	1140

QY 1141 CTGGCGCTGCCACCGAGCGAGCGCAATCAGCTTGTGCGACGAGGTGGAAA 1189  
|||||  
Db 1141 CTGGCGCTGCCACCGAGCGAGCGCAATCAGCTTGTGCGACGAGGTGGAAA 1189  
|||||

Search completed: July 7, 2004, 21:07:43  
Job time : 352.5 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 7, 2004, 20:18:17 ; Search time 2289.2 Seconds  
(without alignments)  
15510.269 Million cell updates/sec

Title: US-10-018-786-3  
Perfect score: 1189  
Sequence: 1 aaattgtcaagggtgatgc.....gcttgcgaccagtgga 1189

Scoring table: IDENTITY\_NJC  
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST:  
1: em\_estba:  
2: em\_esthum:  
3: em\_estin:  
4: em\_estmu:  
5: em\_estov:  
6: em\_estpl:  
7: em\_estro:  
8: em\_htc:  
9: gb\_est1:  
10: gb\_est2:  
11: gb\_htc:  
12: gb\_est3:  
13: gb\_est4:  
14: gb\_est5:  
15: em\_estfun:  
16: em\_estom:  
17: em\_gss\_hum:  
18: em\_gss\_inv:  
19: em\_gss\_pln:  
20: em\_gss\_vrt:  
21: em\_gss\_fun:  
22: em\_gss\_mam:  
23: em\_gss\_mus:  
24: em\_gss\_pro:  
25: em\_gss\_rod:  
26: em\_gss\_phg:  
27: em\_gss\_vrl:

28: gb\_gss1:\*  
29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	145.6	12.2	1259	28	B2555432 pacs1-60
c 2	104.6	8.8	899	28	B2578201 msh2_5747
c 3	103.8	8.7	707	28	BH614295 1C22/2B8
c 4	103.8	8.7	707	28	BH614489 2B8 Conti
c 5	100.6	8.5	1336	28	B2554711 pacs1-60
c 6	97.4	8.2	1361	28	B2575790 msh2_4625
c 7	79.6	6.7	673	28	B2548588 pacs1-60
c 8	76.4	6.4	446	13	C07126 C07126 Rat
c 9	69.6	5.9	787	28	BH795460 04B01LL10
c 10	66.6	5.6	1129	28	B2558715 pgs8401_4
c 11	66.4	5.6	541	28	A2303060 GSSB-u200
c 12	62.6	5.3	1446	28	BH614290 1C22/2B8
c 13	62.6	5.3	1446	28	BH614482 2B8 Conti
c 14	57.6	4.8	414	14	CF143739 UI-HF-BPO
c 15	55.4	4.7	525	29	AL053013 Drosophila
c 16	55	4.6	925	29	AL053013 Drosophila
c 17	51.8	4.4	540	28	B2894800 Hg3_0183
c 18	48.6	4.1	762	29	AL171639 Tetraodon
c 19	48.2	4.1	746	28	B2561680 pacs2-164
c 20	48.2	4.1	935	29	AL066051 Drosophila
c 21	47.8	4.0	790	13	BW263485 BW263485
c 22	47.8	4.0	834	10	BF267459 HV_Cea001
c 23	47.6	4.0	1201	13	BX356664 BX356664
c 24	47.4	4.0	538	29	CC663357 OGWHM17TV
c 25	47.4	4.0	783	28	BZ641592 OGCE132TC
c 26	47.4	4.0	827	29	CC717827 OGJAV26TV
c 27	47	4.0	629	13	BW222739 BW222739
c 28	46.8	3.9	548	6	AL814243 Triticum
c 29	46.8	3.9	567	6	AL815397 Triticum
c 30	46.8	3.9	661	12	B1960553 HVSMEN002
c 31	46.8	3.9	1100	29	AL106955 Drosophila
c 32	46.6	3.9	600	28	BZ896342 NarP8_018
c 33	46.6	3.9	696	14	CA289460 SCAGFL800
c 34	46.6	3.9	766	13	BW101474 BW101474
c 35	46.4	3.9	622	12	BM441961 Eban01_SQ
c 36	46.4	3.9	1100	29	AL106855 Drosophila
c 37	46.2	3.9	580	14	CB878329 HP06C09T
c 38	46.2	3.9	601	28	AY079696 AY079696
c 39	46	3.9	844	29	AL056652 Drosophila
c 40	46	3.9	877	29	CG347796 OGWLH75TV
c 41	45.8	3.8	529	14	CD232008 SSI_31_B0
c 42	45.6	3.8	745	29	CC846671 NDL_10701
c 43	45.6	3.8	906	28	CC122201 NDL_78K22
c 44	45.4	3.8	654	13	CA040006 ssalshc50
c 45	45.4	3.8	742	14	CF774722 UI-D-GC1-